

AD-A185 487

STRONG CONSISTENCY OF M-ESTIMATES FOR THE LINEAR MODEL 1/1

(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS

X R CHEN ET AL JUL 87 TR-87-24 AFOSR-TR-87-0971

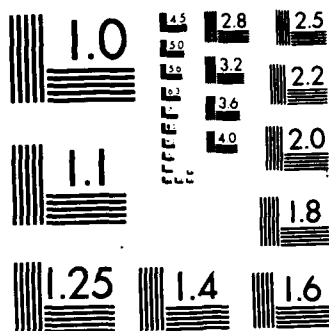
UNCLASSIFIED

F49620-85-C-0008

F/G 12/3

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Unclassified

DTIC FILE COPY (2)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A185 487

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR. 87-0971	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Strong consistency of M-estimates for the linear model		5. TYPE OF REPORT & PERIOD COVERED <i>Journal Technical</i> - July 1987
AUTHOR(s) X. R. Chen and Yeuhua Wu		6. PERFORMING ORG. REPORT NUMBER 87-23
PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis Fifth Floor Thackeray Hall University of Pittsburgh, Pittsburgh, PA 15260		7. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research <i>nm</i> Department of the Air Force Bolling Air Force Base, DC 20332 <i>Bldg. 410</i>		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS <i>61102F 2304 A-5</i>
MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) <i>Scme 335, 11</i>		12. REPORT DATE July 1987
		13. NUMBER OF PAGES 16
		14. SECURITY CLASS. (of this report) Unclassified
		15. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) DTIC ELECTE S OCT 02 1987 D <i>ck D</i>		
18. SUPPLEMENTARY NOTES		
19. Key words and phrases: M-estimate, linear model, strong consistency.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This article defines the M-estimate for the linear model directly from the minimization problem $\sum_{i=1}^n p(Y_i - \alpha - \beta'X_i) = \min.$ (continued)		

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

87 3 24 220

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20.(Continued)

Suppose that $(X_1, Y_1), \dots, (X_n, Y_n), \dots$ are i.i.d. observations of a random vector (X, Y) , where Y is one-dimensional and X may be multi-dimensional. It is shown that the M-estimates $\hat{\alpha}_n, \hat{\beta}_n$ defined in this manner converge with probability one to α_0, β_0 respectively ((α_0, β_0) is the true parameter) as $n \rightarrow \infty$, under very general conditions on the function ρ and the distribution of (X, Y) .

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

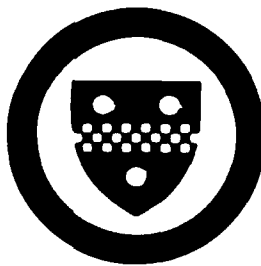
AFOSR-TR. 87-0971

STRONG CONSISTENCY OF M-ESTIMATES
FOR THE LINEAR MODEL *

X. R. Chen and Yeuhua Wu

Center for Multivariate Analysis
University of Pittsburgh

Center for Multivariate Analysis
University of Pittsburgh



STRONG CONSISTENCY OF M-ESTIMATES
FOR THE LINEAR MODEL *

X. R. Chen and Yeuhua Wu
Center for Multivariate Analysis
University of Pittsburgh

July 1987

Technical Report No. 87-24

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260



Accession For	
NTIS	CRA&I <input checked="" type="checkbox"/>
DTIC	TAR <input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution	
Availability Codes	
Dist	Availability Codes
A-1	

* Research sponsored by the Air Force Office of Scientific Research under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

STRONG CONSISTENCY OF M-ESTIMATES FOR THE LINEAR MODEL

X. R. Chen and Yeuhua Wu

ABSTRACT

This article defines the M-estimate for the linear model directly from the minimization problem

$$\sum_{i=1}^n \rho(Y_i - \alpha - \beta'X_i) = \min.$$

Suppose that $(X_1, Y_1), \dots, (X_n, Y_n), \dots$ are i.i.d. observations of a random vector (X, Y) , where Y is one-dimensional and X may be multi-dimensional. It is shown that the M-estimates $\hat{\alpha}_n, \hat{\beta}_n$ defined in this manner converge with probability one to α_0, β_0 respectively ((α_0, β_0') is the true parameter) as $n \rightarrow \infty$, under very general conditions on the function ρ and the distribution of (X, Y) .

AMS 1980 subject classifications: Primary 62J05; secondary 62F12.

Key words and phrases: M-estimate, linear model, strong consistency.

Research sponsored by the Air Force Office of Scientific Research under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

1. INTRODUCTION

Let $(X_1, Y_1), \dots, (X_n, Y_n), \dots$ be i.i.d. observations of a random vector (X, Y) , where Y is one-dimensional and X may be multi-dimensional. Suppose that the regression of Y to X , in some sense, is a linear function $\alpha_0 + \beta_0' X$. It is desired to estimate the unknown parameters α_0, β_0 , by using the observations $(X_1, Y_1), \dots, (X_n, Y_n)$. A much discussed class of estimates is the so-called M-estimate, which takes the solution of the minimization problem

$$\sum_{i=1}^n \rho(Y_i - \alpha - \beta' X_i) = \min \quad (1)$$

as the estimator. Here ρ is a properly selected function defined over $R' = (-\infty, \infty)$.

In literature, the case is often considered in which the X_i 's are supposed to be known constant vectors rather than observations of some random vector. But in many applications, especially in problems of econometrics, it is more practical to assume the random character of the X_i 's.

A common feature of most works dealing with this estimation problem, for example [2], [3] and [6], is to assume that ρ has a continuous derivative ψ everywhere over R' , thus converting the minimization problem (1) to the problem of solving the system of equations

$$\sum_{i=1}^n \psi(Y_i - \alpha - \beta' X_i) = 0, \quad \sum_{i=1}^n X_i \psi(Y_i - \alpha - \beta' X_i) = 0. \quad (2)$$

In order to validate this procedure, one usually makes the assumption that ρ is a convex function. These assumptions seem unduly restrictive for, in some important cases such as $\rho(t) = |t|$ (Minimum L_1 -Norm estimate),

ρ is not everywhere differentiable. Also, the convexity assumption excludes many functions with practical importance. For example, $\rho(t) = \min\{|t|, k\}$ for some constant $k > 0$. In general, any function bounded over R' is not allowed under this assumption.

So it makes much sense to tackle the estimation problem directly starting from the minimization problem (1). Some works (for example, [1], [4], [7]) have been done in this respect for the special case of $\rho(t) = |t|$, but as far as the authors know, no work exists for general ρ up to now.

The purpose of the present article is to study this problem in case that the X_i 's are observations of a random vector. To some extent our method can also be employed to deal with the case in which the X_i 's are known nonrandom vectors, but some additional assumptions will be needed.

2. FORMULATION OF THE RESULTS

In the sequel we shall stick to the notations introduced in Section 1. We shall denote by $(\hat{\alpha}_n, \hat{\beta}_n)$ a Borel-measurable solution of the minimization problem (1).

We shall always impose the following conditions on the function ρ and the random vector (X, Y) :

- a. ρ is continuous everywhere over R^1 .
- b. ρ is nondecreasing over $(0, \infty)$ and nonincreasing over $(-\infty, 0)$.

From b it is seen that $\rho(0) = \min\{\rho(t) : t \in R^1\}$. Without losing generality, we may assume that

$$c. \quad \rho(0) = 0, \quad \rho(t) \geq 0 \quad \text{on } R^1. \quad (3)$$

$$d. \quad E\rho(Y - \alpha - \beta'X) < \infty \quad \text{for any } \alpha \in R^1, \beta \in R^d. \quad (4)$$

- e. For any $x \in R^d$, the function

$$f_x(\theta) \equiv E\{\rho(Y - \theta) | X = x\} \quad (5)$$

attains its minimum uniquely at

$$\theta_x = \alpha_0 + \beta_0'x \quad (6)$$

with α_0, β_0 not depending on x .

For convenience of reference, we shall call the set of conditions $\{a, b, c, d, e\}$ by "Condition (A)".

THEOREM 1. Suppose that the following are true:

- 1. Condition (A).
- 2. $\rho(\infty) = \rho(-\infty) = \infty$.
- 3. If $|\alpha| + \|\beta\| > 0$, then $P(\alpha + \beta'X = 0) < 1$ where $\|\beta\|$ is the Euclidean norm of β .

Then we have as $n \rightarrow \infty$,

$$\hat{\alpha}_n \rightarrow \alpha_0, \quad \hat{\beta}_n \rightarrow \beta_0, \quad \text{a.s.} \quad (7)$$

The following theorem deals with the case that ρ may be bounded.

THEOREM 2. Suppose that the following are true:

1. Condition (A).
2. $\rho(\infty) = \rho(-\infty)$.
3. If $|\alpha| + \|\beta\| > 0$, then $P(\alpha + \beta'X = 0) = 0$.

Then (7) holds true.

Finally, we have the following theorem concerning the convergence rate of $\hat{\alpha}_n$ and $\hat{\beta}_n$.

THEOREM 3. Suppose that

1. The conditions of Theorem 1 or Theorem 2 are true.
2. For any $\alpha \in R^1$ and $\beta \in R^d$, the moment generating function of $\rho(Y - \alpha - \beta'X)$ is finite in some neighborhood of zero (the neighborhood may depend on α, β).

Then for arbitrarily given $\epsilon > 0$, there exists constant $c > 0$ independent of n such that

$$P(|\hat{\alpha}_n - \alpha_0| \geq \epsilon) = O(e^{-cn}), \quad P(\|\hat{\beta}_n - \beta_0\| \geq \epsilon) = O(e^{-cn}) \quad (8)$$

Before entering the details of the proof, we make some remarks about the conditions of the theorems.

1. Condition b seems quite natural from the practical point of view. As for the continuity assumption a, it also seems reasonable. This condition can be weakened to some extent at the expense of a much stronger condition on the distribution of (X, Y) .

2. Condition 3 of Theorem 1, in fact, is a consequence of e. Condition 3 of Theorem 2 holds when X possesses a density.

3. Condition e is closely related to the meaning of the regression. More clearly speaking, the exact meaning of the regression determines the class of functions ρ which can be used in formulating the minimization problem (1). For instance, when $\alpha_0 + \beta_0'x$ is the conditional median of Y given $X = x$ (median regression). We can choose $\rho(t) = |t|$, or any ρ for which $E\{\rho(Y - \theta) | X = x\}$ attains its minimum uniquely at the conditional median. Likewise, when $\alpha_0 + \beta_0'x$ is the conditional expectation of Y given $X = x$, we can choose $\rho(t) = t^2$. An important case is that the conditional distribution of Y given $X = x$ is symmetric and unimodal with center $\alpha_0 + \beta_0'x$. In this case, ρ can be chosen as any even function satisfying conditions a, b, d, and that $\rho(t) > \rho(0)$ when $t \neq 0$.

3. PROOF OF THE THEOREMS

We give the detailed proof of Theorem 1. An easy modification of the argument enables us to prove the remaining two theorems.

The main body of the proof of Theorem 1 is contained in the following two lemmas.

LEMMA 1. Suppose that the conditions of Theorem 1 are satisfied, H is a bounded closed set in R^{d+1} , and $(\alpha_0, \beta'_0) \in H$. Let $(\tilde{\alpha}_n, \tilde{\beta}_n)$ be a Borel measurable solution of the restricted minimization problem

$$\sum_{i=1}^n \rho(Y_i - \alpha - \beta' X_i) = \min \quad (9)$$

with (α, β') being restricted in H . Then as $n \rightarrow \infty$, we have

$$\tilde{\alpha}_n \rightarrow \alpha_0, \quad \tilde{\beta}_n \rightarrow \beta_0, \quad \text{a.s.} \quad (10)$$

Denote by \bar{S}_R the closed ball $\{(\alpha, \beta'): \alpha^2 + \|\beta\|^2 \leq R^2\}$ in R^{d+1} . $(\hat{\alpha}_n, \hat{\beta}'_n)$ is the solution of the unrestricted minimization problem (1) as mentioned earlier.

LEMMA 2. Suppose that the conditions of Theorem 1 are satisfied, then there exists constant R such that with probability one we have $(\hat{\alpha}_n, \hat{\beta}'_n) \in \bar{S}_R$ for n sufficiently large.

It is readily seen that Theorem 1 follows from these two lemmas. Indeed, take $H = \bar{S}_R$, where R is the constant mentioned in Lemma 2. Lemma 2 indicates that

$$P((\hat{\alpha}_n, \hat{\beta}'_n) = (\tilde{\alpha}_n, \tilde{\beta}'_n) \text{ for } n \text{ sufficiently large}) = 1,$$

which in turn entails

$$P(\lim_{n \rightarrow \infty} (\hat{\alpha}_n - \tilde{\alpha}_n) = 0, \lim_{n \rightarrow \infty} (\hat{\beta}_n - \tilde{\beta}_n) = 0) = 1$$

From this and Lemma 1, (7) follows.

Proof of Lemma 1. Without loss of generality, we may assume

$$\alpha_0 = 0, \quad \beta_0 = 0.$$

For any constant $\ell > 0$, define

$$A_\ell = [-\ell, \ell]^{d+1}.$$

Take ℓ large enough such that $\mathbb{H} \subset A_\ell$. Denote the $T = 2^{d+1}$ points $(\pm \ell, \pm \ell, \dots, \pm \ell)$ by $(a_1, b_1'), \dots, (a_T, b_T')$. According to conditions b, c, it can easily be shown that

$$0 \leq \rho(y - \alpha - \beta'x) \leq \sum_{j=1}^T \rho(y - a_j - b_j'x) \quad (11)$$

for any $(x', y) \in \mathbb{R}^{d+1}$ and $(\alpha, \beta') \in A_\ell$. Define

$$Q(\alpha, \beta) = E\rho(Y - \alpha - \beta'x). \quad (12)$$

By (11) and conditions a, d, it follows from the Lebesgue convergence theorem that Q is continuous. By condition e it follows that

$$Q(\alpha, \beta) > Q(0, 0), \quad \text{when} \quad |\alpha| + \|\beta\| > 0. \quad (12)$$

Hence for arbitrarily given $\epsilon > 0$, we have

$$q \equiv \inf\{Q(\alpha, \beta) : (\alpha, \beta') \in \mathbb{H} - A_\epsilon\} - Q(0, 0) > 0. \quad (14)$$

For any constant $M > 0$, denote by $I_M \equiv I_M(X, Y)$ the indicator of the set A_M^C . Choose $\epsilon_1 \in (0, q/4)$ and $M > 0$ large enough such that

$$P\{(X', Y) \in A_M\} > 1 - \epsilon_1, \quad (15)$$

$$E[I_M \rho(Y - \alpha - \beta'x)] < \epsilon_1 \quad \text{for any} \quad (\alpha, \beta') \in \mathbb{H}. \quad (16)$$

The existence of such a constant M follows from condition d and (11).

Write those elements in $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$ which fall into the set A_M , in the order of appearance, as

$$(X_1^*, Y_1^*), \dots, (X_{n'}^*, Y_{n'}^*).$$

Evidently, these variables are conditionally i.i.d. given n' , with the common distribution $(X, Y) | ((X', Y) \in A_M)$. Define the event

$$B_n = \{n' > (1 - 2\epsilon_1)n\}.$$

Then by (15) and the strong law of large numbers, it follows that with probability one B_n occurs for n sufficiently large. Put

$$Q_M(\alpha, \beta) \equiv E[(1 - I_M)\rho(Y - \alpha - \beta'X)].$$

Then Q_M is continuous in α, β . Find $\epsilon_3' > 0$ sufficiently small, such that

$$|Q_M(\tilde{\alpha}, \tilde{\beta}) - Q_M(\alpha, \beta)| \leq \epsilon_1$$

when

$$(\alpha, \beta') \in \mathbb{H}, \quad (\tilde{\alpha}, \tilde{\beta}') \in \mathbb{H}, \quad |\alpha - \tilde{\alpha}| \leq \epsilon_3', \quad \|\beta - \tilde{\beta}\| \leq \epsilon_3'.$$

Put

$$\mathbb{H} \equiv \sup\{|y - a - b'x| : (x', y) \in A_M, (a, b') \in \mathbb{H}\}. \quad (17)$$

Find $\epsilon_2 > 0$ sufficiently small such that

$$\sup\{|\rho(r_2) - \rho(r_1)| : |r_1| \leq \mathbb{H}, |r_2| \leq \mathbb{H}, |r_2 - r_1| \leq \epsilon_2\} < \epsilon_1. \quad (18)$$

Find $\epsilon_3 \in (0, \epsilon_3')$ such that

$$|(a + b'x) - (\tilde{a} + \tilde{b}'x)| \leq \epsilon_2 \quad (19)$$

when

$$(a, b') \in \mathbb{H}, \quad (\tilde{a}, \tilde{b}') \in \mathbb{H}, \quad |a - \tilde{a}| \leq \epsilon_3, \quad \|b - \tilde{b}\| \leq \epsilon_3, \quad \|x\| \leq Md.$$

Choose a finite set $\mathbb{H}_1 = \{(\alpha_i, \beta_i') : i = 1, \dots, m\} \subset \mathbb{H} - A_\epsilon$, such that for any $(\alpha, \beta') \in \mathbb{H} - A_\epsilon$ there exists $(\alpha_i, \beta_i') \in \mathbb{H}_1$ satisfying

$$|\alpha_j - \alpha| \leq \epsilon_3, \quad \|\beta_j - \beta\| \leq \epsilon_3$$

By the strong law of large numbers, with probability one we have

$$\begin{aligned} \frac{1}{n'} \sum_{i=1}^{n'} \rho(Y_i^* - \alpha_j - \beta_j' X_i^*) &> E[\rho(Y - \alpha_j - \beta_j' X)(1 - I_M)] - \epsilon_1 \\ &> E\rho(Y - \alpha_j - \beta_j' X) - 2\epsilon_1 \\ &> Q(0,0) + q - 2\epsilon_1, \quad j = 1, \dots, m \end{aligned} \quad (20)$$

for n sufficiently large. Hence with probability one

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \rho(Y_i - \alpha_j - \beta_j' X_i) &\geq \frac{1}{n} \sum_{i=1}^{n'} \rho(Y_i^* - \alpha_j - \beta_j' X_i^*) \frac{n'}{n} \\ &\geq (1 - 2\epsilon_1)[Q(0,0) + q - 2\epsilon_1] \\ &\equiv Q(0,0) + n', \quad j = 1, \dots, m \end{aligned} \quad (21)$$

for n sufficiently large, where

$$n' = (q - 2\epsilon_1)(1 - 2\epsilon_1) - 2\epsilon_1 Q(0,0).$$

Now choose arbitrarily $(\tilde{\alpha}, \tilde{\beta}') \in \mathcal{H} - A_\epsilon$. Find $(\alpha_j, \beta_j') \in \mathcal{H}_1$ such that $|\alpha_j - \tilde{\alpha}| \leq \epsilon_3$, $\|\beta_j - \tilde{\beta}\| \leq \epsilon_3$. According to (18), (19), (21), we have

$$\begin{aligned} \sum_{i=1}^n \rho(Y_i - \tilde{\alpha} - \tilde{\beta}' X_i) &\geq \sum_{i=1}^{n'} \rho(Y_i^* - \tilde{\alpha} - \tilde{\beta}' X_i^*) \\ &\geq \sum_{i=1}^{n'} \rho(Y_i^* - \alpha_j - \beta_j' X_i^*) \\ &\quad - \sum_{i=1}^{n'} |\rho(Y_i^* - \alpha_j - \beta_j' X_i^*) - \rho(Y_i^* - \tilde{\alpha} - \tilde{\beta}' X_i^*)| \\ &\geq n[Q(0,0) + n'] - n'\epsilon_1 \geq n[Q(0,0) + n], \end{aligned} \quad (22)$$

with probability one for n sufficiently large, where $n = n' - \epsilon_1$.

Choose $\epsilon_1 > 0$ sufficiently small such that $n > 0$. (22) should be understood that it is true simultaneously for all $(\tilde{\alpha}, \tilde{\beta}') \in \mathbb{B} - A_{\epsilon}$ when (21) is true. Therefore, with probability one we have

$$\min\left\{\frac{1}{n} \sum_{i=1}^n \rho(Y_i - \alpha - \beta'X_i) : (\alpha, \beta') \in \mathbb{B} - A_{\epsilon}\right\} \geq Q(0,0) + n \quad (23)$$

for n sufficiently large.

On the other hand, by the strong law of large numbers, with probability one we have

$$\frac{1}{n} \sum_{i=1}^n \rho(Y_i) < Q(0,0) + n \quad (24)$$

for n sufficiently large. From (23), (24), it follows that

$$P(|\hat{\alpha}_n| \leq \epsilon, \|\hat{\beta}_n\| \leq \epsilon, \text{ for } n \text{ sufficiently large}) = 1.$$

This concludes the proof of Lemma 1.

Proof of Lemma 2. Write $S_1 = \{(\alpha, \beta') : |\alpha|^2 + \|\beta\|^2 = 1\}$. According to condition 3 of Theorem 1, one can find $\epsilon > 0$ such that

$$q \equiv \inf\{P(|\alpha + \beta'X| > \epsilon) : (\alpha, \beta') \in S_1\} > 0. \quad (25)$$

Find M sufficiently large, such that $P(X \in A_M) > 1 - q/4$. Put $n = [3(1+dM)]^{-1}$. Choose a finite set $\tilde{S}_1 \subset S_1$ in such a way so that for any $\theta \in S_1$ there exists $\tilde{\theta} \in \tilde{S}_1$ satisfying $\|\theta - \tilde{\theta}\| < n$. By (25) and the strong law of large numbers, with probability one we have

$$\#\{i : 1 \leq i \leq n, |\alpha + \beta'X_i| > \epsilon \text{ for any } (\alpha, \beta') \in \tilde{S}_1\} \geq nq/2 \quad (26)$$

for n sufficiently large, where $\#(A)$ denotes the number of elements of set A .

By the strong law of large numbers, with probability one we have

$$\#\{i: 1 \leq i \leq n, X_i \in A_M\} \geq n(1 - q/4). \quad (27)$$

Choose a constant K such that

$$qK/8 > Q(0,0) + 1$$

where the function Q is defined by (12). According to condition 2 of Theorem 1, one can find h so that $\rho(a) \geq K$ when $|a| \geq h$. Choose a constant R large enough so that

$$\epsilon R/4 \geq h,$$

$$P(|Y| < \epsilon R/4) > 1 - q/8.$$

By the strong law of large numbers, with probability one we have

$$\#\{i: 1 \leq i \leq n, |Y_i| \leq \epsilon R/4\} \geq n(1 - q/8). \quad (28)$$

Now choose arbitrarily a point (α, β') outside \bar{S}_R . We have

$$(\alpha, \beta') = r(\tilde{\alpha}, \tilde{\beta}'): r > R, \quad (\tilde{\alpha}, \tilde{\beta}') \in \tilde{S}_1.$$

Assume that (26)-(28) are true, then

1. If $(\tilde{\alpha}, \tilde{\beta}') \in \tilde{S}_1$, from (26) we have

$$\#\{i: 1 \leq i \leq n, |\alpha + \beta' X_i| \geq R\epsilon\} \geq nq/2. \quad (29)$$

From (28), (29), we obtain

$$\#\{i: 1 \leq i \leq n, |Y_i - \alpha - \beta' X_i| \geq 3R\epsilon/4\} \geq 3nq/8. \quad (30)$$

2. If $(\tilde{\alpha}, \tilde{\beta}') \notin \tilde{S}_1$, then choose $(\alpha^*, \beta^{*'}) \in \tilde{S}_1$ satisfying

$$|\tilde{\alpha} - \alpha^*| \leq \eta, \quad \|\tilde{\beta} - \beta^{*'}\| \leq \eta,$$

when $|\alpha^* + \beta^{*'} X_i| > \epsilon$ and $X_i \in A_M$, we have

$$\begin{aligned} |\tilde{\alpha} + \tilde{\beta}' X_i| &> \epsilon - |(\alpha^* - \tilde{\alpha}) + (\beta^{*'} - \tilde{\beta}') X_i| \\ &\geq \epsilon - |\alpha^* - \tilde{\alpha}| - \|\beta^{*'} - \tilde{\beta}'\| \cdot \|X_i\| \\ &\geq \epsilon - \eta - \eta dM > \epsilon/2. \end{aligned}$$

Hence $|\alpha + \beta'X_i| > R\epsilon/2$. From this, and (26)-(28), we get

$$\#\{i: 1 \leq i \leq n, |Y_i - \alpha - \beta'X_i| > R\epsilon/4\} \geq nq/8. \quad (31)$$

Summarizing these two cases, we see that with probability one,

$$\frac{1}{n} \sum_{i=1}^n \rho(Y_i - \alpha - \beta'X_i) \geq qK/8 > Q(0,0) + 1$$

holds simultaneously for all (α, β') outside \bar{S}_R , when n is large enough.

Since (24) is true for $n = 1$, we see that with probability one

$$\min\left\{\frac{1}{n} \sum_{i=1}^n \rho(Y_i - \alpha - \beta'X_i) : (\alpha, \beta') \notin \bar{S}_R\right\} > \frac{1}{n} \sum_{i=1}^n \rho(Y_i) \quad (32)$$

for n sufficiently large. Therefore

$$P\{(\hat{\alpha}_n, \hat{\beta}'_n) \in S_{\bar{R}} \text{ for } n \text{ sufficiently large}\} = 1 \quad (33)$$

which proves Lemma 2.

Proof of Theorem 2. Without loss of generality, assume $\alpha_0 = 0$, $\beta_0 = 0$. No change is needed in the proof of Lemma 1. For a proof of Lemma 2 under the conditions of Theorem 2, first note that $Q(0,0) < L(\infty)$ by conditions a, b, e and condition 2 of Theorem 2. From condition 3 of Theorem 3, it is readily seen that for any $a < 1$ there exists $\epsilon > 0$ such that

$$\inf\{P(|\alpha + \beta'X| > \epsilon) : (\alpha, \beta') \in S_1\} > a.$$

Starting from these two facts and employing the argument used earlier, it can be shown that there exists constant R such that

$$P\left\{\frac{1}{n} \sum_{i=1}^n \rho(Y_i - \alpha - \beta'X_i) > c \text{ for } (\alpha, \beta') \notin \bar{S}_R\right. \\ \left. \text{simultaneously, when } n \text{ large enough}\right\} = 1 \quad (34)$$

where $Q(0,0) < c < L(\infty)$. From (34) we obtain (32), hence (33).

Proof of Theorem 3. The proof follows from the following two lemmas:

LEMMA 1'. Under the conditions of Theorem 3 for arbitrarily given $\epsilon > 0$, there exists constant $c > 0$ such that

$$P(|\tilde{\alpha}_n - \alpha_0| \geq \epsilon) = O(e^{-cn})$$

$$P(|\tilde{\beta}_n - \beta_0| > \epsilon) = O(e^{-cn})$$

where $\tilde{\alpha}_n, \tilde{\beta}_n$ are the same as in Lemma 1.

LEMMA 2'. Under the conditions of Theorem 3, there exist constants $R > 0$ and $c > 0$ such that

$$P\{(\hat{\alpha}_n - \alpha_0, \hat{\beta}_n' - \beta_0') \notin \overline{S}_R\} = O(e^{-cn}).$$

These lemmas can be proved by employing the argument used in proving Lemma 1 and Lemma 2, with the help of the following fact (see [5], p.288): Suppose that ξ_1, ξ_2, \dots is a sequence of i.i.d. random variables, $E\xi_1 = 0$ and $E \exp(t\xi_1) < \infty$ for $|t| < \delta$, $\delta > 0$, then for arbitrarily given $\epsilon > 0$ there exists constant $c > 0$ such that

$$P(|\sum_{i=1}^n \xi_i / n| \geq \epsilon) = O(e^{-cn}).$$

REFERENCES

- [1] BLOOMFIELD, P. and STIEGER, W.L. (1983). *Least Absolute Deviations*. Birkhäuser.
- [2] HUBER, P.J. (1973). Robust regression *Ann. Statist.* 1, 799-821.
- [3] MARONNA, R.A. and YOHAI V.J. (1981). Asymptotic behavior of general M-estimates for regression and scale with random carriers. *Z. Wahrsch. Verw. Gebiete* band 58, 7-20.
- [4] OBERHOFER, W. (1982). The consistency of nonlinear regression minimizing the L_1 -norm. *Ann. Statist.* 10, 316-319.
- [5] PETROV, V.V. (1975). *Sums of Independent Random Variables*. Springer-Verlag, New York.
- [6] YOHAI, V.J. and MARONNA, R.A. (1979). Asymptotic behavior of m-estimators for the linear model. *Ann. Statist.* 7, 258-268.
- [7] WU, Yeuhua (1987). Strong consistency and exponential rate of the "minimum L_1 -norm" estimates in linear regression models. Tech. Report. No. 87-18, Center for Multivariate Analysis, Univ. of Pittsburgh.

END

DATE
FILMED

DEC.

1987